

# 階層型強化学習 RGoal をベースにした 記号推論と強化学習の統合の構想

全脳アーキテクチャシンポジウム

2019-06-26

産業技術総合研究所

人工知能研究センター

一杉裕志

参考:「BESOM(ビーソム)ブログ 記号推論と強化学習を  
統合した脳型汎用人工知能アーキテクチャの構想」

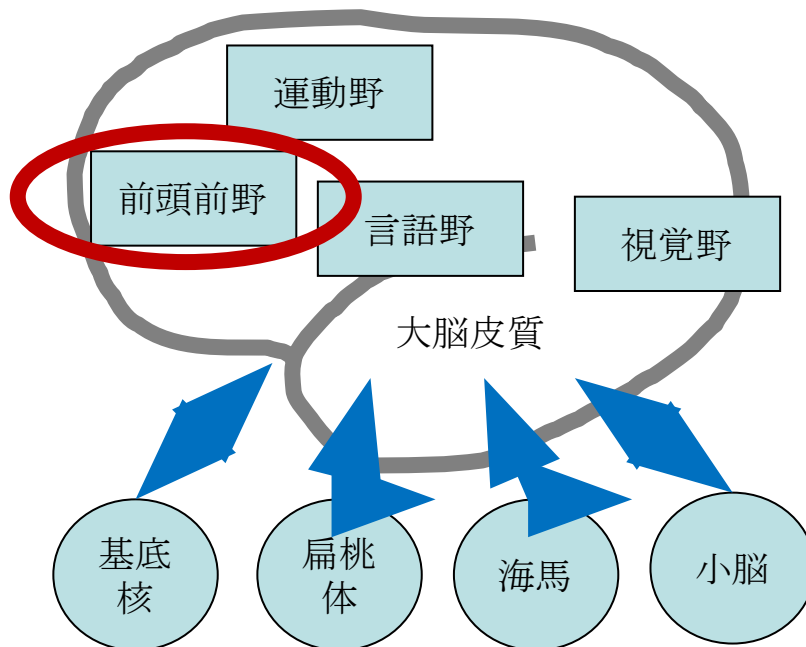
<http://besom1.blog85.fc2.com/blog-entry-238.html>

# なぜAGIができていないのか

- 最大の問題はAGI実現に取り組む研究者不足。
- 研究者を増やすためには、多くの人にAGIの実現可能性を確信させる成果を出すことが、まず必要。

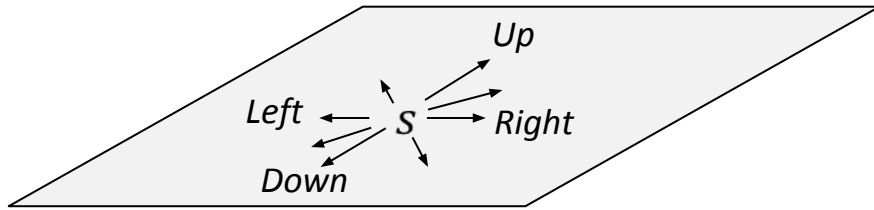
そこで、ヒトの知能の中核である前頭前野周辺のモデル構築に取り組む中。

→ 記号推論と強化学習の統合を目指す。

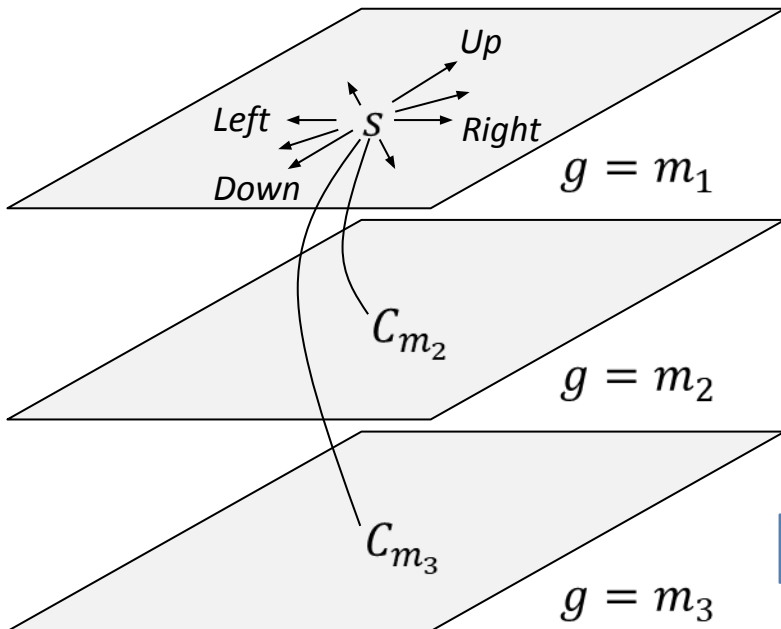


# RGoal は再帰的なサブルーチン呼び出しができる階層型強化学習

[一杉 et al. 2018 汎用人工知能研究会]



状態行動空間の  
拡張



サブルーチンの再利用により学習を加速。

$$\begin{aligned} \mathcal{S} &= \{(0,0), (0,1), \dots\} \\ \mathcal{A} &= \{Up, Down, Right, Left, \dots\} \end{aligned}$$

もともとの状態  $s$  とサブゴール  $g$  の組  $\tilde{s} = (s, g)$  を拡張された状態と見なす。

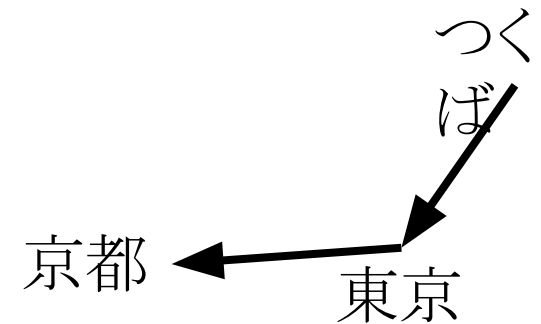
$$\begin{aligned} \tilde{\mathcal{S}} &= \mathcal{S} \times \mathcal{M} \\ \tilde{\mathcal{A}} &= \mathcal{A} \cup \mathcal{C}_{\mathcal{M}} \\ \mathcal{M} &= \{m_1, m_2, \dots\} \subseteq \mathcal{S} \\ \mathcal{C}_{\mathcal{M}} &= \{C_{m_1}, C_{m_2}, \dots\} \end{aligned}$$

RGoal は拡張された状態行動空間上のMDPを解くアルゴリズム

$$Q(s, g, a) \leftarrow Q(s, g, a) + \alpha(r + Q(s', g', a') - Q(s, g, a) + V_g(g'))$$

# RGoal の「思考モード」は演繹推論

- エージェントが獲得済みの知識を組み合わせて、未知の問題を解くための機構。
  - 一種のモデルベース強化学習。
  - 時間を抽象化した脳内シミュレーションを行う。  
例:「京都に行くには東京駅に出て新幹線に乗ればよい」



既知の公理を組み合わせて新たな定理を導く  
**演繹推論の機構**と見なせる。

さらに、一階述語論理に相当する表現力を持たせることを検討中

# 記号推論と機械学習の統合のカギ： ベイジアンネット

例：パターンマッチを行うベイジアンネットの回路  
(同様の回路で記号推論の1ステップも実現可能)

(X,X,1,Y,Y,4)

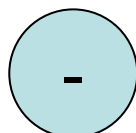
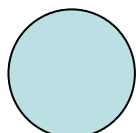
パターン内の変数への値の割り当て

パターン P

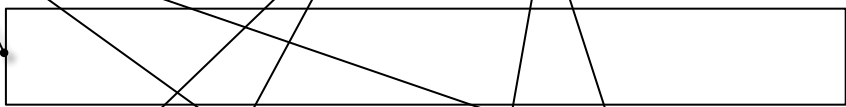
X

Y

Z



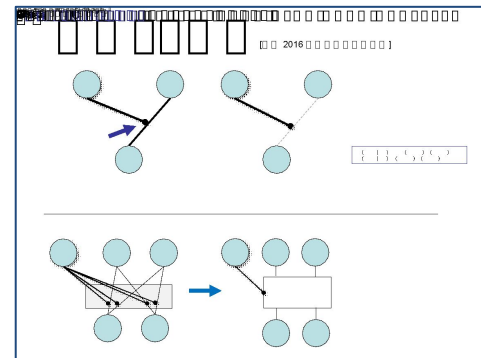
パターンに応じて  
ゲートの開閉を制御



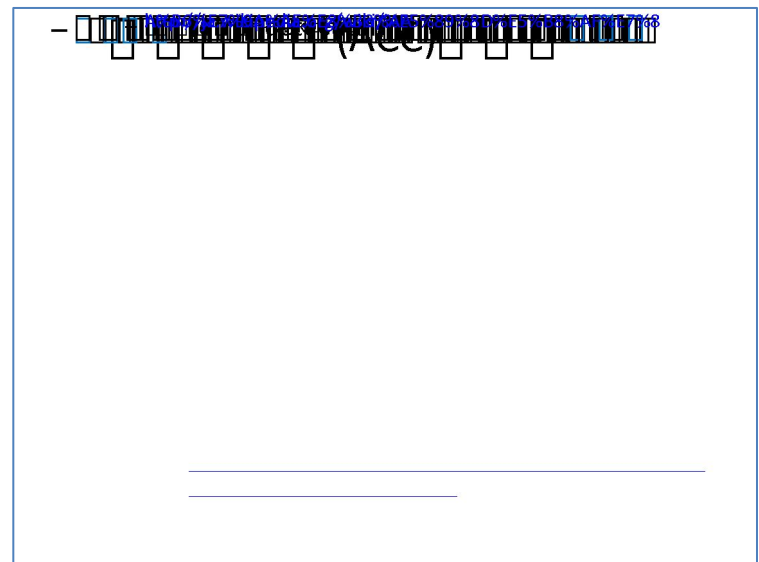
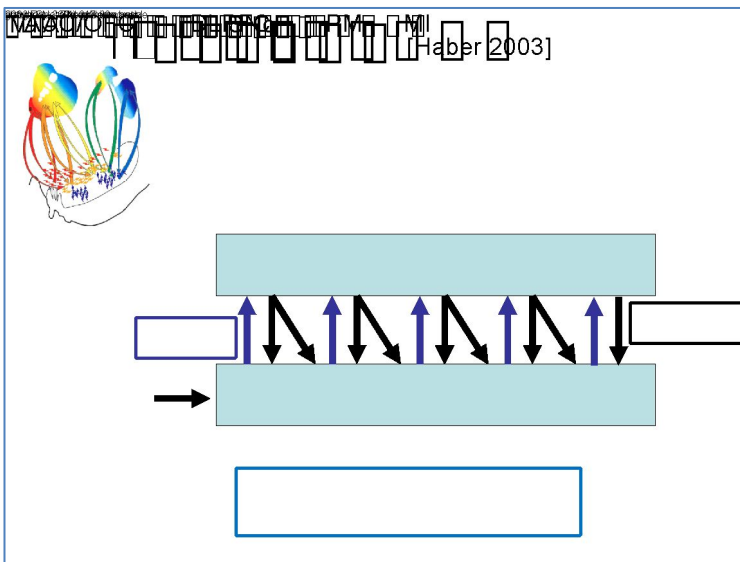
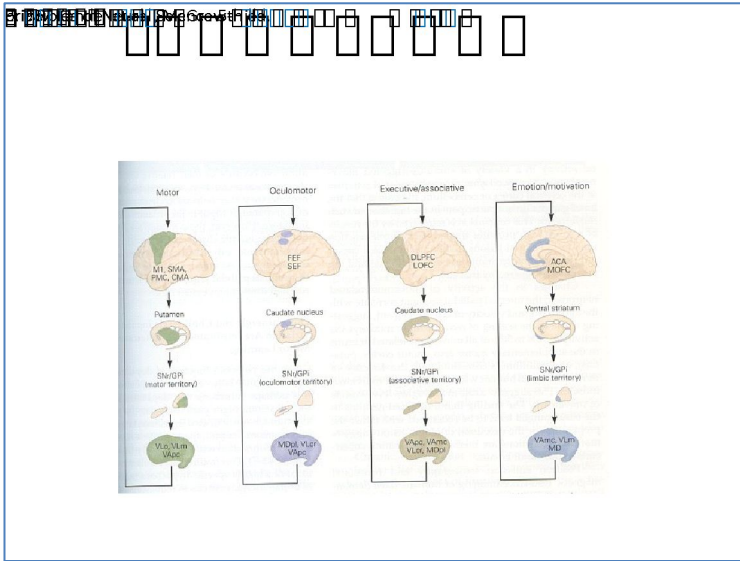
下の層に値を与えると、上の層にマッチするパターンと変数の値が推論される。

入力ベクトル

ゲートノード群



# 参考にすべき神経科学的知見



# まとめ

- **記号推論と強化学習を統合**することでAGI研究の飛躍的加速を目指す。
- 統合のカギとなる技術が**再帰的な階層型強化学習とベイジアンネット**。
- **脳のマクロな解剖学的構造**はとてもよいヒントになる。
- いずれも世界でほとんど手が付けられておらず、出し抜くチャンス。

# 第27回 全脳アーキテクチャ勉強会

- テーマ「確率的グラフィカルモデルと脳」
- 7月19日(金)東大 本郷 鉄門記念講堂
- 講演者:

**IBM東京基礎研究所**

**恐神貴行 氏**

「動的ボルツマンマシンと脳(仮)」

**NTT コミュニケーション科学基礎研究所**

**石畠正和 氏**

「ZDD による指数関数的な計算時間の高速化(仮)」